

Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/JP05/008495

International filing date: 10 May 2005 (10.05.2005)

Document type: Certified copy of priority document

Document details: Country/Office: JP
Number: 2004-140117
Filing date: 10 May 2004 (10.05.2004)

Date of receipt at the International Bureau: 24 June 2005 (24.06.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)



World Intellectual Property Organization (WIPO) - Geneva, Switzerland
Organisation Mondiale de la Propriété Intellectuelle (OMPI) - Genève, Suisse

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日 Date of Application: 2004年 5月10日

出願番号 Application Number: 特願2004-140117

パリ条約による外国への出願に用いる優先権の主張の基礎となる出願の国コードと出願番号

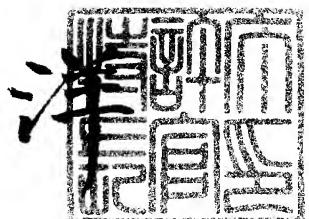
The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is

出願人 Applicant(s): 日本電信電話株式会社

2005年 6月 8日

特許庁長官
Commissioner,
Japan Patent Office

小川



【書類名】 特許願
【整理番号】 NTTH157619
【提出日】 平成16年 5月10日
【あて先】 特許庁長官殿
【国際特許分類】 G10L
【発明者】
【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】 大室 伸
【発明者】
【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】 森 岳至
【発明者】
【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】 日和▲崎▼ 祐介
【発明者】
【住所又は居所】 東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
【氏名】 片岡 章俊
【特許出願人】
【識別番号】 000004226
【氏名又は名称】 日本電信電話株式会社
【代理人】
【識別番号】 100066153
【弁理士】
【氏名又は名称】 草野 卓
【選任した代理人】
【識別番号】 100100642
【弁理士】
【氏名又は名称】 稲垣 稔
【手数料の表示】
【予納台帳番号】 002897
【納付金額】 16,000円
【提出物件の目録】
【物件名】 特許請求の範囲 1
【物件名】 明細書 1
【物件名】 図面 1
【物件名】 要約書 1
【包括委任状番号】 9806848

【書類名】特許請求の範囲

【請求項 1】

音声・音楽などの音響信号をパケット通信網を介して通信する方法において、送信側では、

音響信号をフレームと呼ばれる一定時間ごとに区切り、

各フレーム内の音響信号に対応する少くともピッチ周期を音響特徴量として求め、

遅延量制御情報として指定されるフレーム数だけ、フレーム番号が異なる音響信号と音響特徴量を同一のパケットに組み込んで送信し、

受信側では、

受信したパケットを受信バッファに蓄え、

上記受信バッファからフレーム番号の順にパケットを取り出し、パケット内の音響信号を順次連結して出力し、

出力しようとするフレームの音響信号に対応するパケットが受信できなかった（以下紛失したという）場合に、

上記紛失パケットの音響信号に対応する音響特徴量を上記受信バッファ内のパケットから求め、

上記紛失パケットの音響信号フレーム（以下紛失フレームという）と直近したフレームの音響信号から、上記音響特徴量に含まれるピッチ周期に対応する長さの波形を切り出し、

その切り出した波形を上記ピッチ周期で繰り返し並べた信号を用いて上記紛失パケットの音響信号を生成することを特徴とする音響信号のパケット通信方法。

【請求項 2】

音声・音楽などの音響信号をフレームと呼ばれる一定時間ごとに区切り、

各フレーム内の音響信号に対応する少くともピッチ周期を音響特徴量として求め、

上記音響特徴量を、遅延量制御情報として指定されるフレーム数だけ、フレーム番号が異なるフレームの音響信号と同一のパケットに組み込んで送信することを特徴とする音響信号のパケット送信方法。

【請求項 3】

上記各フレーム内の音響信号のパワも上記音響特徴量として求めることを特徴とする請求項 2 記載の音響信号パケット送信方法。

【請求項 4】

送信先側から受信したパケットから相手側バッファ残量を取り出し、上記相手側バッファ残量に応じて上記遅延量制御情報を生成し、

上記受信したパケットを蓄え音響信号を再生出力するため用いる受信バッファの残量を上記パケットに組み込むことを特徴とする請求項 2 又は 3 記載の音響信号のパケット送信方法。

【請求項 5】

受信したパケットを受信バッファに蓄え、

上記受信バッファからフレーム番号の順にパケットを取り出し、パケット内の音響信号を順次連結して出力し、

出力しようとするフレームの音響信号に対応するパケットが受信できなかった（以下紛失したという）場合に、

上記紛失パケット（以下紛失フレームという）の音響信号に対応する音響特徴量を上記受信バッファ内のパケットから求め、

上記紛失フレームと直近した正常に受信されたフレームの音響信号から、上記音響特徴量に含まれるピッチ周期に対応する長さの波形を切り出し、

その切り出した波形を上記ピッチ周期で繰り返し並べ、その並べた信号を用いて上記紛失パケットの音響信号を生成することを特徴とする音響パケットの受信方法。

【請求項 6】

上記直近したフレームの音響信号は、上記紛失フレームの直前のフレームの出力音響信

号であって切り出した波形を前向きに並べた信号（以下第1音響信号という）と上記紛失フレームより後で最も近いフレームのパケットから得られる音響信号であって、切り出した波形を後ろ向きに並べた信号（以下第2音響信号という）とであり、

上記第1音響信号と第2音響信号とにそれぞれ重みを乗じた後、加算して上記紛失パケットの音響信号を生成することを特徴とする請求項5記載の音響信号のパケット受信方法。

【請求項7】

上記紛失フレームが連續した複数であり、上記生成する紛失パケットの音響信号のフレーム番号と上記直前のフレーム番号及び上記後で最も近いフレームの番号とのそれとの差（第1差及び第2差という）を求め、

上記第1音響信号及び上記第2音響信号とにそれぞれ乗じる第1重み及び第2重みの割り合いを、上記第2差及び第1差と対応させることを特徴とする請求項6記載の音響信号のパケット受信方法。

【請求項8】

上記紛失フレームの音響信号に対応する音響特徴量を求めることができない場合、

上記波形を切り出す音響信号が上記紛失フレームの前の直近フレームの信号であれば、上記紛失フレームの直前のフレームの出力音響信号からピッチ周期を求め、そのピッチ周期と対応する長さで上記波形の切り出しを行い、

上記波形を切り出す音響信号が上記紛失フレームの後の直近フレームの信号であれば、上記紛失フレームの直後のフレームの音響特徴量に含まれるピッチ周期と対応する長さで上記波形の切り出しを行うことを特徴とする請求項5～7のいずれかに記載の音響信号のパケット受信方法。

【請求項9】

上記音響特徴量に含まれるパワと同等になるように上記生成した音響信号の振幅を補正することを特徴とする請求項5～8のいずれかに記載の音響信号のパケット受信方法。

【請求項10】

上記受信パケットに組み込まれている遅延量制御情報を参照して、上記紛失フレームの音響信号に対応する音響特徴量を求めることが特徴とする請求項5～9のいずれかに記載の音響信号のパケット受信方法。

【請求項11】

上記ピッチ周期に対応する長さは、上記ピッチ周期より少し長く、上記繰り返し並べる際に重なり部分を重み付加算することを特徴とする請求項5～10のいずれかに記載の音響信号のパケット受信方法。

【請求項12】

入力音響信号をフレームと呼ばれる一定時間ごとに区切ってパケット構成部でパケットを構成してパケット通信網に送信する装置において、

各フレーム内音響信号に対応するピッチ周期を表す音響特徴量をフレームと呼ばれる一定時間ごとに区切って計算する音響特徴量計算部と、

遅延量制御情報として指定されるフレーム数だけ対応音響信号に対し遅延させて上記パケット構成部へ供給する遅延手段と

を具備する音響信号パケット送信装置。

【請求項13】

受信したパケットを受信バッファに蓄え、受信パケットを順に取り出し、パケット内の音響信号をフレーム番号順に連結して音響信号を出力する装置において、

上記受信パケットが入力され、パケットの紛失を検出する紛失検出部と、

紛失が検出されたパケットの音響信号と対応するピッチ周期を表す音響特徴量を上記受信バッファに蓄えられているパケットの中から得る受信バッファ探索部と、

上記紛失パケットの音響信号のフレームの直前のフレームと対応する、上記出力される音響信号と、上記ピッチ周期が入力され、上記直前の出力音響信号から、上記ピッチ周期に対応する長さの波形を切り出して、上記ピッチ周期で繰り返し並べて上記紛失パケット

の音響信号を外挿生成する紛失信号生成部と
を具備する音響信号パケット受信装置。

【請求項 1 4】

請求項 2～4 のいずれかに記載した音響信号のパケット送信方法の各過程をコンピュータに実行させるためのプログラム。

【請求項 1 5】

請求項 5～11 のいずれかに記載した音響信号のパケット受信方法の各過程をコンピュータに実行させるためのプログラム。

【請求項 1 6】

請求項 1 4 又は 1 5 のいずれかに記載したプログラムを記録したコンピュータ読み取り可能な記録媒体。

【書類名】明細書

【発明の名称】音響信号のパケット通信方法、送信方法、受信方法、これらの装置及びプログラム、その記録媒体

【技術分野】

【0001】

この発明は、デジタル化された音声・音楽などの音響信号をインターネットをはじめとするパケット通信網を介して送信する際に、特にパケット紛失対策をした通信方法、送信方法、受信方法、これらの装置及びプログラム、その記録媒体に関する。

【背景技術】

【0002】

音声信号をボイスオーバ (Voice over) IP (インターネットプロトコル) 技術を利用して送信するサービスが普及しつつある。図1に示すように入力端子11よりの音声信号を音声信号送信部12で音声パケットに変換してIP網をはじめとするパケット通信網13によって音声信号受信部14へ送信し、音声信号受信部14により音声信号を再生して出力端子15へ出力する。これをリアルタイム通信する場合、通信網13の状態によっては通信網の途中においてパケットロス (紛失) が生じ、それによって再生音声が途切れるといった品質劣化が問題となっている。特に、インターネットなどのベストエフォートと呼ばれる通信サービスの場合には、パケットロスを許容しているため通信網の混雑時に特にこの問題が顕著である。

【0003】

そこで、音声信号をパケット通信網で通信する場合には、パケットロスコンシールメントと呼ばれる手法を用いて、パケットが通信路の途中で消失あるいは通信路の遅延によって制限時間内に受信側に届かなかった場合 (以下パケットロス (紛失) の場合) に、消失または届かなかったパケット (以下ロスパケット又は紛失パケットという) に対応する区間の音声信号を受信側で推定して補償する方法が用いられる。図2は、図1における音声信号送信部12の一般的な構成例である。入力音声は入力バッファ21に蓄えられ、音声パケット化部22で音声信号をフレームと呼ばれる一定の時間ごとに区切って音声パケットを生成し、パケット送出部23よりパケット通信網に音声パケットを送出する。1フレームの時間長は一般には、10ミリ秒から20ミリ秒程度が多い。

【0004】

図3は、図1における音声信号受信部14の一般的な構成例である。パケット通信網からパケット受信部31で受信した音声パケットは、ゆらぎ吸収バッファとも呼ばれる受信バッファ32に蓄えられた後、正しくパケットが受信されたフレームについては、音声パケット復号部33で音声信号に復号され、パケットロスしたフレームについては、紛失信号生成部34でパケットロスコンシールメント処理を行って音声信号が出力される。パケットロスコンシールメントの処理に、ピッチ周期 (音声の基本周波数に相当する時間軸上の長さ) の情報を利用する場合には、出力音声信号を出力音声バッファ35に蓄え、ピッチ抽出部36においてピッチ分析し、得られたピッチ周期の値を紛失信号生成部34に供給する。紛失信号生成部34よりの生成された信号は切替スイッチ37を通じて出力端子15へ出力され、パケットロスがない場合は音声パケット復号部33よりの復号音声信号が切替スイッチ37を通じて出力端子15へ出力される。なお、双方向で音声通信を行う通信端末は、各端末に送信部と受信部の両方を具備する。パケットロスコンシールメントの代表的な方法としては、非特許文献1に示す方法がよく知られている。非特許文献1に示す方法では、音声のピッチ周期をパケットロスコンシールメントに利用している。

【0005】

図4に、非特許文献1に示す方法でも用いられている、一般的なパケットロスコンシールメントの手法を示す。この図では、受信側の現在のフレームをフレームnとして、フレームnに対応するパケットがロス (紛失) した場合の処理方法を示すものである。過去のフレーム (フレームn-1まで) の音声信号は正しくパケットが受信されて復号されているか、パケットロスがあった場合にはパケットロスコンシールメントの手法によって既に

音声信号が生成されているものとする。フレームnでは音声パケットが受信できていないため、直前のフレームn-1の最後のサンプル点から、1ピッチ周期分の区間3Aの音声信号波形を切り出し、フレームnの区間に切り出した1ピッチ波形を順にフレームの長さになるまで並べる（区間3B～3D）。

【0006】

このように直前フレームの1ピッチ波形を並べる処理によってパケットロスとなったフレームの波形を生成すると、何も処理しないで0値でフレームnの全サンプル点を埋めるのに比べて、自然な音質で再生することが可能である。

なお、このとき、1ピッチ波形を単純に並べた場合、接続点で波形が不連続となって、ツツツツといった耳障りな音が生じことがある。その場合には、図5に示すような手法で接続点での不連続を防ぐ。図5では、説明をわかりやすくするために、フレームnの区間を、切り出し波形ごとに段をずらして表示している。まず、フレームn-1の最後のサンプル点から、1ピッチ周期よりも少し長い、例えばピッチ長をLとすると4分の5L（ $5 \times L / 4$ ）の区間4Aの波形を切り出し、1ピッチ長ずつずらした4B, 4C, 4Dの位置に切り出した波形をそれぞれ並べる。このとき、切り出した波形は1ピッチ長よりも長いため、重なり区間4AB, 4BC, 4CDができる。これらの重なり部分は、例えば図6に示す三角窓関数をかけて重ね合わせるという方法がとられる。図6は、横軸が時刻、縦軸が重みの値を示しており、t1が重なり区間の始点、t2が重なり区間の終点を示し、例えば図5における重なり区間4BCでは、重なり区間4BCの区間4B側の波形に重み関数W1を乗じ、区間4C側の波形に重み関数W2を乗じて加算することにより、区間BとCの両切り出し波形をなめらかにつなげることができる。このような重ね合わせの詳細についても、非特許文献1に示す方法の標準勧告に記載されている。

【0007】

パケットロスの発生する通信環境で非特許文献1に示す方法を利用した場合の音声品質は一般に良好と言われているが、パケットロスが音声の子音と母音の境界付近で発生した場合に、耳障りな雑音が発生することがある（課題1）。また、連続した複数のフレームにおいてパケットロスが発生（バースト的なロス（紛失）という）したり、1フレームのみのパケットロスでも、フレーム長が長い音声符号化方式を利用した場合には、耳障りなブザー音が発生したり、不自然な音声が再生されるといった問題がある（課題2）。なお、「連続した複数のフレームにおいてパケットロスが発生」「フレーム長が長い」とは、例えばパケットロスが原因で40ミリ秒～60ミリ秒程度以上連続して正常な音声受信ができない状態が該当する。

【0008】

前記課題1の原因是、非特許文献1に示す方法がロスフレームの音声を生成するための手法として、直前のフレームの音声波形と基本的に特性が同じ波形を作成していることに起因する。即ち、子音と母音の境界付近の母音に近いフレームでフレームロスが発生すると、ロスしたフレームは母音区間であるにもかかわらず、子音と同じ特性の音声波形が生成される。また、母音から無音や子音へ変化する時刻でも同様の雑音が発生する。

前記課題2の原因是、仮にパケットロスの発生した区間が子音と母音の境界付近でなかったとしても、非特許文献1に示す方法を利用して生成されたパケットロスフレームにおける音声をさらに（自己回帰的に）利用して、連続した後続のフレームロス区間においても同じ特性の音声波形が生成されるため、結果として、40ミリ秒～60ミリ秒以上の長い時間にわたって、同じ特性の音声波形が連続して再生されることに起因する。実際の音声はピッチ周期やパワが微小に変化しており、連続して同じ特性の音が再生されると、それは音声とは異なった音に聞こえる。

【非特許文献1】ITU-T Recommendation G.711 Appendix I, “A high quality low-complexity algorithm for packet loss concealment with G.711”, pp. 1-18, 1999

9

【発明の開示】

【発明が解決しようとする課題】

【0009】

パケットロスの発生する通信環境において、非特許文献1に示す方法のパケットロスコンシールメント手法を用いることは、再生音声の品質劣化を少なくする点で一定の効果がある。しかしながら、子音と母音の境界付近のフレームでパケットロスが発生した場合には、耳障りな雑音を十分に防ぐことができない、また連続した複数のフレームでパケットロスが発生した場合の再生品質が不自然である。この発明はこれらの問題を解決し、より安定した品質で音響通信のできる方法、その装置、プログラム、その記録媒体を提供することを目的とする。

【課題を解決するための手段】

【0010】

この発明によれば、送信側で、各フレーム内の音響信号に対応する少くともピッチ周期を音響特徴量として求め、遅延量制御情報として指定されるフレーム数だけ、フレーム番号が異なる音響信号と音響特徴量を同一のパケットに組み込んで送信し、

受信側で、出力しようとするフレームの音響信号に対応するパケットが紛失した場合は、紛失パケットの音響信号に対応する音響特徴量を、受信バッファ内のパケットから求め、上記紛失パケットの音響信号のフレーム（以下紛失フレームという）と直近したフレームの音響信号から、上記求めた音響特徴量に含まれるピッチ周期に対応する長さの波形を切り出し、その切り出した波形を上記ピッチ周期で繰り返し並べ、この並べた信号を用いて上記紛失パケットの音響信号を生成する。

【発明の効果】

【0011】

この発明によれば、各フレームの音響信号のピッチ周期を、その音響信号とは別のパケットで送信しているから、あるフレームの音響信号がパケットロスにより失なわれても、その音響信号のピッチ周期はパケットロスすることなく受信され、このピッチ周期と対応する長さで音響信号波形を切り出し、並べて、紛失音響信号を生成しているため、1フレーム単位のランダムなパケットロスの場合でも、連続して複数フレームのパケットロスが発生した場合でも、パケットロスが発生しない場合に近い音響品質で音響信号が再生されるため、安定した音響通信が実現される。

【0012】

従ってパケットロスの発生頻度の高いパケット通信網でリアルタイムの安定した音響通信ができ、また、パケット通信網はコストを抑えるために、ある程度のパケットロスを許容するように設計するのが一般的であり、この発明の利用によって、回線自体のパケットロス率が低い高品質ネットワークが不要となるため、ネットワークコスト低減にも効果がある。

【発明を実施するための最良の形態】

【0013】

この発明は、コンピュータ本体とコンピュータプログラムとして実行することが可能であるし、デジタルシグナルプロセッサや専用LSIに実装して実現することも可能である。この発明は音声・音楽などの音響信号に適用できるが、以下では、図1に示したVoice over IPによる音声通信システムにこの発明を適用した場合について説明する。音声信号送信装置100（図1中の送信部12と対応）の機能構成例を図7に、音声信号受信装置200（図1中の受信部14と対応）の機能構成例を図8にそれぞれ示し、音声信号送信装置100の処理手順の例を図9に、音声信号受信装置200の処理手順を図10にそれぞれ示す。

送信側

送信装置100において、入力音声信号は入力バッファ111に蓄えられ、音声信号をフレームと呼ばれる一定の時間ごとに区切って、つまりフレーム分割して（ステップS1）音声波形符号化部112に送る。1フレームの時間長は一般には、10ミリ秒から20ミリ秒程度とすることが多い。

【0014】

音声波形符号化部112では、例えば、ITU-Tの標準符号化方式であるG.711(μ-law PCMまたはA-law PCM方式)を適用して入力音声信号を音声符号に変換する(ステップS2)。なお、G.711以外の標準方式や、非標準の音声符号化方式など、符号化方式は任意の方式を用いてよい。音声符号はパケット構成部113に送られる。

音声特徴量計算部114では、入力バッファ111に蓄えられた音声信号を用いて、当該フレームにおけるその音声信号の音声特徴量を計算する(ステップS3)。音声特徴量とは、ピッチ周期(音声の基本周波数に相当)、パワなどを指し、これらの特徴量のいずれかのみを利用することもあるし、全部を利用することもある。ピッチ周期はピッチ部114aで例えば、音声信号波形または音声信号波形にスペクトル包絡の逆特性を持つフィルタをかけた信号の自己相関係数を計算することによって得られる(ステップS3a)。パワは、パワ部114bでフレーム内の音声信号の全サンプルの二乗和により計算する(ステップS3b)。この二乗和をフレーム長のサンプル数(音声信号のサンプリング周波数が8kHzでフレーム長が20ミリ秒の場合は160)で除して平方根をとった値(1サンプルあたりのパワの平方根)をパワを表す特徴量として利用してもよい。

【0015】

音声特徴量符号化部115では、前記音声特徴量を決められたビット数(またはバイト数)で表すことができるよう量子化した後、符号に変換する(ステップS4)。例えば、ピッチ周期の場合、音声のサンプリング周波数が8kHzサンプリングで、ピッチ周期の探索範囲が40サンプル～120サンプルの場合には、符号として0～80を割り当てることができ、7ビットで符号化することができる。探索範囲が20サンプル～160サンプルの場合は符号として0～140を割り当てるので8ビットである。パワの場合は、前記1サンプルあたりのパワの平方根にG.711方式を適用すれば8ビットで符号化することができる。

【0016】

符号化された音声特徴量は、シフトバッファ116に送られる。シフトバッファ116では、あらかじめ指定された数のフレームにわたって音声特徴量の符号を保持し、端子117よりの後述する遅延量制御情報で指定されるフレーム数だけ前、つまり過去のフレームの音声信号の音声特徴量の符号(補助情報ともいう)をパケット構成部113に送る。例えば、現在のフレームをn、遅延量制御情報が3とすると、フレームn-3の音声信号の音声特徴量符号がパケット構成部113に送られる。フレームn-3で生成された補助情報がシフトバッファ116で3フレーム遅延されてパケット構成部113へ送られる(ステップS5)。

【0017】

端子118よりの後述するバッファ残量がバッファ残量符号化部119で符号化され(ステップS6)、そのバッファ残量符号もパケット構成部113に送られる。パケット構成部113では、前記音声信号波形を符号化した符号と、音声特徴量の符号と、遅延量制御情報と、バッファ残量符号を用いてパケットを構成する(ステップS7)。なお、遅延量制御情報とバッファ残量符号はパケットに組み入れない場合もある。この点については後述する。

パケット送出部121は、パケット構成部113で作成されたパケットの情報を受け取り、音声パケットとしてパケット通信網に送出する(ステップS8)。

【0018】

パケット構成部113においてパケットを構成する際の、パケットの構成例を図11に示す。一般にパケットはヘッダ領域41とデータ領域42からなり、ヘッダ領域41の中でも必須領域41aとオプション領域41bがある。パケットの宛先やフレーム番号などのタイムスタンプはヘッダ領域41に格納される。音声信号波形を符号化した符号データはデータ領域42に格納される。音声特徴量の符号は、波形データに比べてサイズ(バイト数)が少ないので、ヘッダ領域41のオプション領域HOP41bに格納する方法と、データ領域42の先頭D1または末尾DEに格納する方法のいずれでもよい。オプション

領域HOPに格納する方法が、この発明を適用しない方法との互換性を維持しやすいという長所がある。また、パケットのヘッダ領域41は4バイトの倍数とするのが一般的である。音声特徴量として、ピッチ周期とパワをそれぞれ1バイトずつで量子化して送る場合（7ビットの場合は1バイト=8ビットに切り上げる）、ピッチ周期とパワ情報の合計2バイトに、遅延量制御情報とバッファ残量符号を各1バイトで表現して合計4バイトとなる。これらを例えればオプション領域HOP41bの4つの領域OP1, OP2, OP3及びOP4に各1バイトずつ格納する。遅延量制御情報を補助情報（音声特徴量符号）とともに送るのは、補助情報がどのフレームに対応する補助情報であるかを、受信側で知るためにあり、フレームnを基準とした相対的なタイムスタンプとみることができる。また、バッファ残量符号を補助情報とともに送るのは、相手側が自分宛にパケットを送るときに、補助情報をフレームnと相対的に何フレームずれで送ってもらうのが適当であるかを伝えるためである。

【0019】

補助情報がフレームnと相対的に何フレームずれたものであるかをフレーム毎にダイナミックに変動させない場合は、遅延量制御情報やバッファ残量符号を補助情報とともに送ることは不要であり、図7中のバッファ残量符号化部119、図8中のバッファ残量復号部216、受信バッファ残量判定部218、図9中のステップS6は省略される。あらかじめ決められた相対的なフレームのずれを送信側と受信側で事前に（例えは呼の接続時のネゴシエーションで）知っていればよいからである。その場合には、遅延量制御情報とバッファ残量符号を各1バイトで送る必要がなく、ピッチ周期とパワ情報の合計2バイトだけで済む。ヘッダ領域が4バイト単位であるので、補助情報を2フレーム分、例えは、フレームn-3とフレームn-6の各音声信号に対する各補助情報をフレームnのパケットにつけて送ることができる。それだけ音声信号符号とその補助情報のいずれもがパケットロスとなるおそれがなくなる。またこの遅延量制御情報を一定にする場合は、音声信号符号と、その補助情報が共にパケットロスにならないように、つまりバースト的ロスに耐え、かつリアルタイム性が失われない程度に、60ミリ秒、80ミリ秒あるいは100ミリ秒程度と対応するフレーム数とする。

受信側

図8の音声信号受信装置200においてはパケット受信部211は、パケット通信網から音声パケットを受信し（ステップS21）、受信バッファ212に蓄積する（ステップS22）。受信バッファ212はゆらぎ吸収バッファとも呼ばれる。前述したように、音声パケットには、音声信号波形を符号化した符号と、ピッチ、パワ等の符号の補助情報、遅延量制御情報、バッファ残量符号が含まれている。補助情報がフレームnと相対的に何フレームずれたものであるかをフレーム毎にダイナミックに変動させない場合は、遅延量制御情報とバッファ残量符号は不要であるが、以下は遅延量制御情報とバッファ残量符号がともにパケットに組み込まれていることを前提にして説明する。

【0020】

受信した音声パケットに含まれる音声信号波形を符号化した符号は、音声パケット復号部213に送られ、音声信号波形に復号される（ステップS24）。パケットロスが発生していないフレームでは音声パケット復号部213の出力信号が切替スイッチ214を通じて出力端子215に再生音声として出力される（ステップS36）。

受信した音声パケットに含まれるバッファ残量符号は、バッファ残量復号部216において、補助情報を何フレームずらしてパケットにつけるかを指定する遅延量制御情報として、図7中の端子117、つまりシフトバッファ116とパケット構成部113に送られる。相手側バッファ残量と遅延量制御情報の関係については後述する。

【0021】

受信した音声パケットに含まれる遅延量制御情報は、紛失処理制御部217で利用される。紛失処理制御部217における処理は後で詳細に述べる。

受信バッファ残量判定部218は、受信バッファ212に蓄積されているパケットのフレーム数を検出する。例えは、フレームnが音声パケット復号部213で復号されている

ときに、受信バッファ212にフレームn+1, n+2, n+3の音声信号符号のパケットが蓄積されている状態であれば、バッファ残量、つまり蓄積パケットのフレーム数は3である。このバッファ残量は、図7中の端子118、つまりバッファ残量符号化部119に送られる。なお、図7に示す送信装置100と、図8に示す受信装置200が相互に連携するこの通信方法は、双方向音声通信を行い、各端末が送信装置100及び受信装置200の両方を具備していることが前提である。

【0022】

紛失検出部219はパケットロス（紛失）を検出する（ステップS23）。例えばパケット受信部211で受信されたパケットはそのパケット番号、つまりフレーム番号の順に受信バッファ212に格納され、つまり受信したパケットのフレーム番号がその直前に受信したパケットのフレーム番号より3つ多ければ、直前に受信したパケットの格納位置から2つのパケットを格納する位置をあけてその受信パケットを格納する。受信バッファ212からはその格納位置の順にパケットが読み出されるが、読み出される際に読み出す格納位置にパケットがない場合には、その読み出し動作の直前に、パケットロス（紛失）が発生したと紛失検出部219が判定し、切替スイッチ214を紛失処理制御部217の出力側に切替える。紛失処理制御部217におけるパケットロスコンシールメント制御処理の詳細を説明する。

【0023】

フレームnにおいて、パケットロスが発生したとする。パケットロスが発生した場合は、受信バッファ探索部221において受信バッファ212に蓄積された受信パケットを探索し、フレームn+1以降で受信されているパケットのうち、紛失フレームnに時間の近い（タイムスタンプが直近の）パケットを探索する（ステップS25）。この探索の結果フレームn+iのパケットが得られたとする。例えば、フレームn+1がロス（紛失）していないときはi=1、フレームn+1、フレームn+2とも連続してロスしてフレームn+3がロスでない場合はi=3といった具合である。フレームn+iのパケットに含まれる音声信号波形を符号化した符号は、先読み音声波形復号部222で復号され（ステップS26）、フレームn+iの音声信号波形が得られる。さらに受信バッファ探索部221において受信バッファ212に蓄積されたパケットの中から、紛失フレームnの音声信号に対応する補助情報が付加されたパケットを探索する（ステップS27）。

【0024】

前述したように、補助情報がフレームnと相対的に何フレームずれたパケットにつけられているかという遅延量制御情報（相対的なタイムスタンプ）がパケットに付加されている場合には、その遅延量制御情報をもとにフレームnの音声信号と対応する補助情報が付加されたパケットを探索することができる。例えば、フレームn+3のパケットに遅延量制御情報として3が付加されていれば、それはフレームnの音声信号と対応する補助情報である。補助情報が付加されるパケットはフレームnと相対的に何フレームずれたものであるかをフレーム毎にダイナミックに変動させない場合は、遅延量制御情報を送信側でパケットに付加する必要がなく、その場合は、あらかじめ決められた遅延量を前提として紛失フレームnの音声信号と対応する補助符号が付加されたパケットを探索する。例えば、あらかじめ決められた遅延量が5のとき、紛失フレームnの音声信号対応補助情報は、フレームn+5のパケットに付加されている。前述したように1フレーム分のパケットに2フレーム分の補助情報を付加している場合は、探索の結果、紛失フレームnの音声信号対応補助情報が付加されたパケットは受信バッファ212内に2つ見つかることがあるので、その場合はいずれかを利用すればよい。

【0025】

このパケット探索で見つかれば（ステップS28）、探索された紛失フレームnの音声信号に対応する補助情報は、音声特徴量復号部223において紛失フレームnの音声信号のピッチ情報とパワ情報に復号され（ステップS29）、紛失信号生成部224に送られる。

一方、非特許文献1に示す方法と同様に、出力端子215に出力される出力音声信号は

出力音声バッファ 225 に蓄えられ (ステップ S36) 、ステップ S28 でパケット検索により見つからなければ出力音声バッファ 225 の出力音声信号はピッチ抽出部 226 においてピッチ周期の分析が行われる (ステップ S30) 。ピッチ抽出部 226 において抽出されるピッチは、紛失フレームの直前フレーム $n-1$ の音声信号に対応するピッチである。直前フレーム $n-1$ の音声信号に対応するピッチは、紛失信号生成部 224 に送られる。

【0026】

紛失信号生成部 224 の動作イメージを図 12 に、機能構成例を図 13 に示す。フレーム n を現在のフレームとし、フレーム n に対応するパケットがロス (紛失) したものとする。非特許文献 1 に示す方法は、パケットロスの発生した直前のフレーム $n-1$ の波形をピッチ単位で繰り返すもので、ピッチ周期、パフを含む信号の特性は直前のフレームと同一であった。この実施形態では図 13 に示すように、直前のフレーム $n-1$ の信号からピッチ周期と対応する信号波形を切り出し、この信号波形を、非特許文献 1 に示す方法と同様に、前向き波形外挿部 51 においてピッチ周期単位で波形を繰り返して 1 フレーム長の波形 10-1 を作成する (ステップ S31) 。このとき、非特許文献 1 に示す方法では、ピッチ周期として直前のフレーム $n-1$ における音声信号のピッチ長を用いているが、この発明では、補助情報を復号して得られた紛失フレーム n の音声信号のピッチ長を用いる。補助情報を組み込んだパケットがロスした場合など、ステップ S28 でパケットを見つけることができず補助情報が得られていない場合は、ステップ S30 で得られた直前のフレーム $n-1$ の音声信号のピッチ長で代用する。

【0027】

また、この実施形態ではフレーム $n+1$ 以降で、受信バッファ 212 内に正常に受信できたフレーム $n+i$ の音声信号、つまり、フレーム $n+1$ のパケットが受信できていればフレーム $n+1$ の信号、フレーム $n+1$ のパケットも連続してロスして、フレーム $n+2$ のパケットが受信できた場合にはフレーム $n+2$ の信号からピッチ周期と対応する信号波形を切り出し、後ろ向き波形外挿部 52 においてピッチ周期単位で波形を繰り返して 1 フレーム長の波形 10-2 を作成する (ステップ S32) 。このとき、時間軸上で未来の波形を使って外挿するため、波形の繰り返し方向が時間軸の負の向きであることに注意する。また後ろ向き波形外挿部 52 で使用するピッチ周期は前向き波形外挿部 51 と同様に、補助情報を復号して得られた紛失フレーム n の音声信号のピッチ長とするが、補助情報を組み込んだパケットもロスした場合など、補助情報が得られていない場合は、フレーム $n-1$ の音声信号のピッチ長で代用するか、または、フレーム $n+1$ の音声信号対応の補助情報が得られていれば、フレーム $n+1$ の音声信号のピッチ長で代用してもよい。

【0028】

つまり紛失フレーム n の音声信号を生成しようとする場合に波形外挿の際の切り出しピッチ長は例えば図 14 に示すように、紛失フレーム n の音声信号に対する補助情報が格納されたパケットがロスでなければ (ステップ S41) 、その補助情報を復号して用い (ステップ S42) 、その補助情報が付加されたパケットがロスであれば、外挿が前向きかを調べ (ステップ S43) 、前向きであれば、直前のフレーム $n-1$ の出力音声信号からピッチ長を分析して、そのピッチ長を用い (ステップ S44) 、外挿が前向きでなければ、直後のフレーム $n+1$ の音声信号に対する補助情報が付加されたパケットを探索し、これがロスでなければ、ステップ S42 に移ってその補助情報を復号してピッチ長を用いる。なおこのフレーム $n+1$ の音声信号の補助情報も得られなければ、ステップ S44 に移る。前記前向き波形外挿部 51 と、後ろ向き波形外挿部 52 のそれぞれ処理イメージを図 12A と図 12B に示す。前向き波形外挿では、フレーム $n-1$ の音声信号から切り出した 1 ピッチ長の波形 10A1 を区間 10B1, 10C1, 10D1 のように繰り返して波形 10-1 を作成する。後ろ向き波形外挿では、フレーム $n+1$ の音声信号から切り出した 1 ピッチ長の波形 10A2 を、区間 10B2, 10C2, 10D2 のように繰り返して波形 10-2 を作成する。図 15B に、フレーム n 、フレーム $n+1$ が連続してパケットロスになり、フレーム $n+2$ が正常に受信できた場合の後ろ向き波形外挿のイメージを示す

。この場合は、フレーム $n+2$ の音声信号から 1 ピッチ波形 1 3 A を切り出し、これをフレーム $n+1$ のフレーム $n+2$ 側からフレーム n のフレーム $n-1$ 側に向って区間 1 3 B, 1 3 C, …, 1 3 F と繰り返すことによってフレーム n の外挿波形 1 0-2 を得る。

【0029】

なお、非特許文献 1 に示す方法の説明において図 5 で述べたように、単純に 1 ピッチ長に切り出した波形を並べると、接続点で不連続音が出るため、前向き波形外挿では図 5 と同様に重なり区間を設けて窓掛け加算 (Overlap add, OLA) を行うのがよい。同様に後ろ向き波形外挿においても図 1 6 に示すように、窓掛け加算 (OLA) を、図 5 の時間軸を逆転した方法で実現できる。

波形 1 0-1 と波形 1 0-2 には、図 1 3 に示すように乗算部 5 3 と乗算部 5 4 においてそれぞれ重み W_1 と W_2 が乗算され、加算部 5 5 において加算され波形 y_n となる (ステップ S 3 4)。重み W_1 と W_2 はそれぞれ $1/2$ にする簡略化された方法と、波形 1 0-1 と波形 1 0-2 の「信頼度」に基づいて重み決定部 5 6 でそれぞれ重み W_1 と W_2 を決定する (ステップ S 3 4) 方法がある。「信頼度」に基づくとは、過去最後に正常受信できたフレームと現在のフレーム n とのフレーム番号差と、受信バッファ 2 1 2 内から探索して得た現在よりも後の時刻 (タイムスタンプ) の受信パケットのフレームと現在のフレーム n とのフレーム番号差とのうち、番号差の小さいほうを信頼度が高いとみなすことで、バースト的なパケットロスの場合に有用である。具体例を以下にいくつか挙げる。

【0030】

例 1：直前のフレーム $n-1$ ではパケットロスが発生しておらず、フレーム n でパケットロスが発生、直後のフレーム $n+1$ ではパケットロスが発生しなかった場合は、前向き外挿波形 1 0-1 と後ろ向き外挿波形 1 0-2 の「信頼度」は同じとみなされるので、重みはそれぞれ $1/2$ とする。

例 2：直前フレーム $n-1$ ではパケットロスが発生しておらず、フレーム n でパケットロスが発生、直後フレーム $n+1$ もパケットロス、更に次のフレーム $n+2$ ではパケットロスが発生しなかった場合は、前向き外挿波形 1 0-1 と後ろ向き外挿波形 1 0-2 の「信頼度」は直前のフレーム $n-1$ から前向き波形外挿によって得られた波形 1 0-1 のほうが高いと推測される。したがって、前向き外挿波形 1 0-1 に乘ずる重みは $2/3$ 、後ろ向き外挿波形 1 0-2 に乘ずる重みは $1/3$ とする。これは正常にパケットが受信された前、後の直近フレーム $n-1$ と $n+2$ から音声信号を生成しようとするフレーム n までの「距離」が $1:2$ であるため、信頼度を $2:1$ とみなす。

【0031】

例 3：フレーム $n-3$ が正常受信、フレーム $n-2$ 、フレーム $n-1$ が連続パケットロスの後、フレーム n でパケットロス、フレーム $n+1$ ではパケットロスが発生しなかった場合は、前向き外挿波形 1 0-1 より後ろ向き外挿波形 1 0-2 のほうが信頼度が高いと推測される。前向き外挿波形 1 0-1 に乘ずる重みは $1/4$ 、後ろ向き外挿波形 1 0-2 に乘ずる重みは $3/4$ とする。これは正常にパケットが受信された前、後直近のフレームから音声信号を生成しようとするフレーム n までの「距離」が $3:1$ であるため、信頼度を $1:3$ とみなす。従って図 1 3 中の重み決定部 5 6 では例えれば正常にパケットが受信された前及び後のフレーム n_B 及び n_L と音声信号を生成しようとするフレーム n_N とのそれぞれの差 Δn_1 及び Δn_2 が差回路 5 6 a 及び 5 6 b で計算され、これらの差 Δn_1 及び Δn_2 が和回路 5 6 c で加算され、この加算値により差 Δn_2 を除算回路 5 6 d で割算して前向き外挿波形 1 0-1 に対する重み $W_1 = \Delta n_2 / (\Delta n_1 + \Delta n_2)$ を求め、また前記加算値により差 Δn_1 を除算回路 5 6 e で割算して後ろ向き外挿波形 1 0-2 に対する重み $W_2 = \Delta n_1 / (\Delta n_1 + \Delta n_2)$ を求める。

【0032】

前記例では、前向き外挿波形 1 0-1 の全サンプル点、後ろ向き外挿波形 1 0-2 の全サンプル点に対して、フレーム内で一律の重みを乗じたが、正常にパケットが受信されたフレームからの「距離」をサンプル点ごとに決定して、サンプル点毎に信頼度に応じた重みを設定してもよい。

図13中のパワ補正部57は、加算波形 y_n のパワを、フレーム n の音声信号の補助情報を復号して得られたパワ情報によって補正し(ステップS35)、フレーム n の音声信号 x_n として出力端子215に出力する(ステップS36)。パワを補正するとは加算波形 y_n のパワが補助情報が示すパワと同等になるように波形 y_n の振幅を増減することである。ここで、同等とは、完全にパワが一致するか、または聴感上の影響を考慮して出力する音声信号 x_n のパワのほうが多少低くなる程度にすることをいう。

バッファ残量と遅延量制御情報

前述の相手側バッファ残量と遅延量制御情報の関係について説明する。

【0033】

遅延量(=何フレームずらして補助情報を送るかのずれ量)は、大きければ大きいほど、バースト的なロスに強くできるが、受信側で補助情報をを利用してパケットロスコンシールメント(紛失パケットの音声信号の生成)の処理をするには、補助情報が到着するまで音声の再生ができないため、通話の遅延が増大する。逆に、遅延量が小さいと、通話の遅延も小さくできる反面、単独の(ランダムな)パケットロスには耐えられるものの、バースト的なロスの場合には、補助情報も失われ、補助情報を音声波形の符号と並行して送る効果が得られないというトレードオフが生じる。そのような場合に最適なのは、通信相手側の受信バッファにその時点で何フレーム分のパケットが蓄積されているかという情報を相手側から受け取って、相手側受信バッファにおけるパケット蓄積フレーム数に相当する遅延量(補助情報を何フレームずらすかの量)を設定することである。

【0034】

前述したように、パケットの到着時間のゆらぎを吸収するために受信バッファには一定量のパケットが蓄積されるようにするのが一般的であり、パケット通信網の通信状態があまりよくない場合には、相手側は受信バッファに蓄積するパケットの数を増やして、パケットの到着ゆらぎによって音切れが生じないようになるのが一般的であり、もともとその受信バッファ蓄積パケット数分の遅延が生じている。従って、相手側受信バッファのパケット蓄積フレーム数(バッファ残量)に相当する値を遅延量として設定して補助情報を送信すれば、新たな遅延を増加させることなく、補助情報送信の効果を最大にすることができる。なお、相手側受信バッファのパケット蓄積フレーム数以下の遅延量であれば、新たな遅延を増加させることができないので、想定されるバースト的なパケットロスの特性に応じて、相手側受信バッファのパケット蓄積フレーム数以下の遅延量に設定してもよく、相手側受信バッファのバッファ残量が時々刻々変動することを考慮すれば、相手側バッファ残量よりも1~2フレーム程度少ない遅延量に設定する方法も考えられる。

【0035】

音響特徴量としてはピッチ周期のみでもよい。この場合は図7中のパワ部114b、図9中のステップS3b、図10中のステップS35、図13中のパワ補正部57は省略される。また送信側で音響信号波形を符号化することなく、例えば入力されたPCM音響信号波形をそのままパケットに組み込んでもよい。この場合は、図7中の音声波形符号化部112、図8中の音声パケット復号部213、先読み音声波形復号部222、図9中のステップS2、図10中のステップS24、ステップS26は省略される。更に音響特徴量についても、これらは通常ディジタル演算により求められるから、その計算結果が例えば8ビットで得られるようにすれば、特に音響特徴量も符号化する必要がなく、図7中の音声特徴量符号化部115、図8中の音声特徴量復号部223、図9中のステップS4、図10中のステップS29は省略される。

【0036】

紛失信号の生成において、前後のパケットが正常に受信され、フレーム n の1パケットのみが紛失(ロス)した場合、前向き外挿波形10-1のみあるいは後ろ向き外挿波形のみを用いてもよい。またバースト的パケットロスの場合は、正常に受信された直後のパケットの音響信号としては前向き外挿波形のみを、正常に受信された直前のパケットの音響信号に対しては、後ろ向き外挿波形のみを用い、これらの中間の紛失パケットに対しては、前向き外挿波形と後ろ向き外挿波形との重み付加算としてもよい。

図7に示した音響信号パケット送信装置100、図8に示した音響信号パケット受信装置200はそれぞれコンピュータにより機能させてもよい。その場合は、図9に示したパケット送信方法の処理手順又は図10に示したパケット受信方法の処理手順の各過程をコンピュータに実行させるための音響信号パケット送信プログラム又は音響信号パケット受信プログラムをコンピュータにCD-ROM、磁気ディスク装置、半導体記憶装置などの記録媒体からインストールし、あるいは通信回線を介してダウンロードしてそのプログラムをコンピュータに実行させればよい。

【図面の簡単な説明】

【0037】

【図1】音声信号を音声パケットに変換し、パケット通信網によって通信する概念図。

【図2】図1中の音声信号送信部12の一般的な機能構成例を示すブロック図。

【図3】図1中の音声信号受信部14の一般的な機能構成例を示すブロック図。

【図4】一般的なパケットロスコンシールメントの手法による紛失信号の生成を説明するための波形図。

【図5】図4においてピッチ波形を繰り返して波形を作成する際に不連続音が生じないようにするための重ね合わせ手法（オーバーラップアンド処理）を説明するための波形図。

【図6】オーバーラップアンド処理における三角窓関数の例を示す図。

【図7】この発明による音響信号パケット送信装置の機能構成例を示すブロック図。

【図8】この発明による音響信号パケット受信装置の機能構成例を示すブロック図。

【図9】この発明による音響信号パケット送信方法の処理手順の例を示す流れ図。

【図10】この発明による音響信号パケット受信方法の処理手順の例を示す流れ図。

【図11】パケットの構成例を示す図。

【図12】図8中の紛失信号生成部224の動作の一例を説明するための波形図。

【図13】図8中の紛失信号生成部224の具体的機能構成例を示すブロック図。

【図14】紛失フレームの音響信号の特徴量の選択手順の例を示す流れ図。

【図15】連続してパケットロスが発生した場合の後ろ向き波形外挿処理を説明するための波形図。

【図16】後ろ向き波形外挿のオーバーラップアンド処理を説明するための波形図。

【書類名】 図面

【図 1】

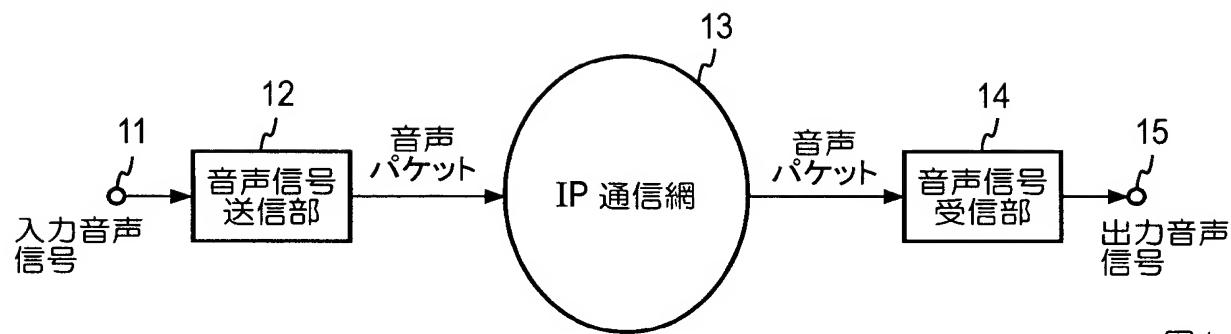


図1

【図 2】

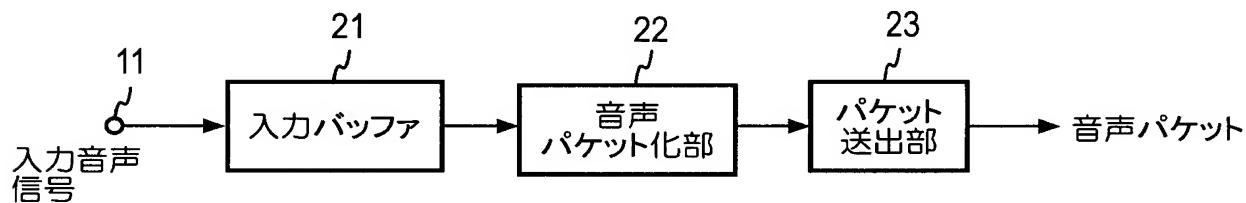


図2

【図 3】

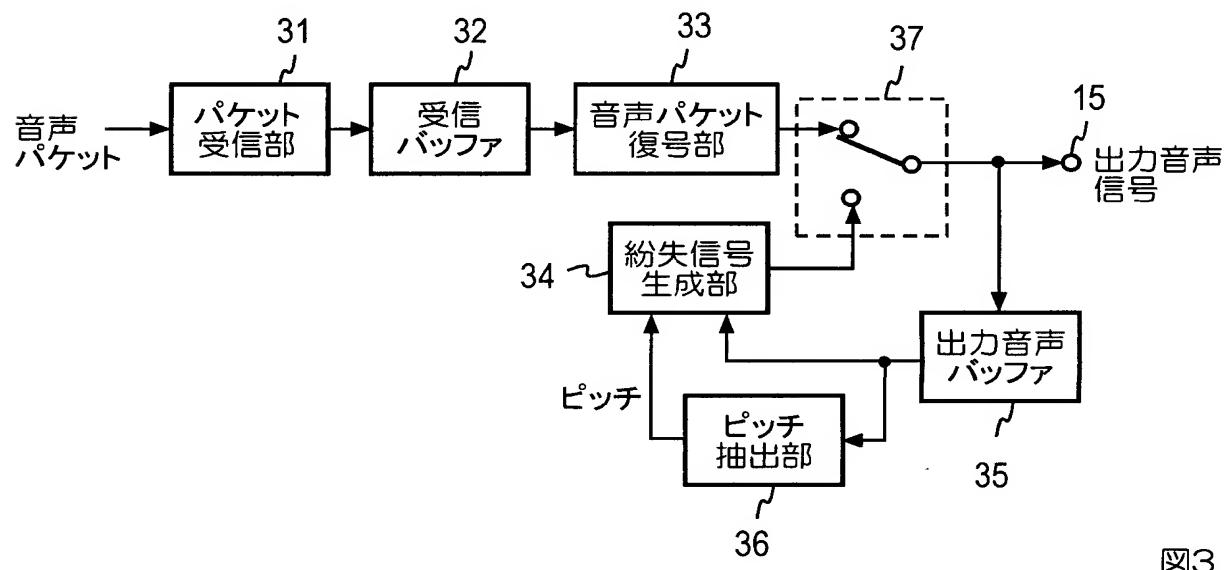


図3

【図 4】

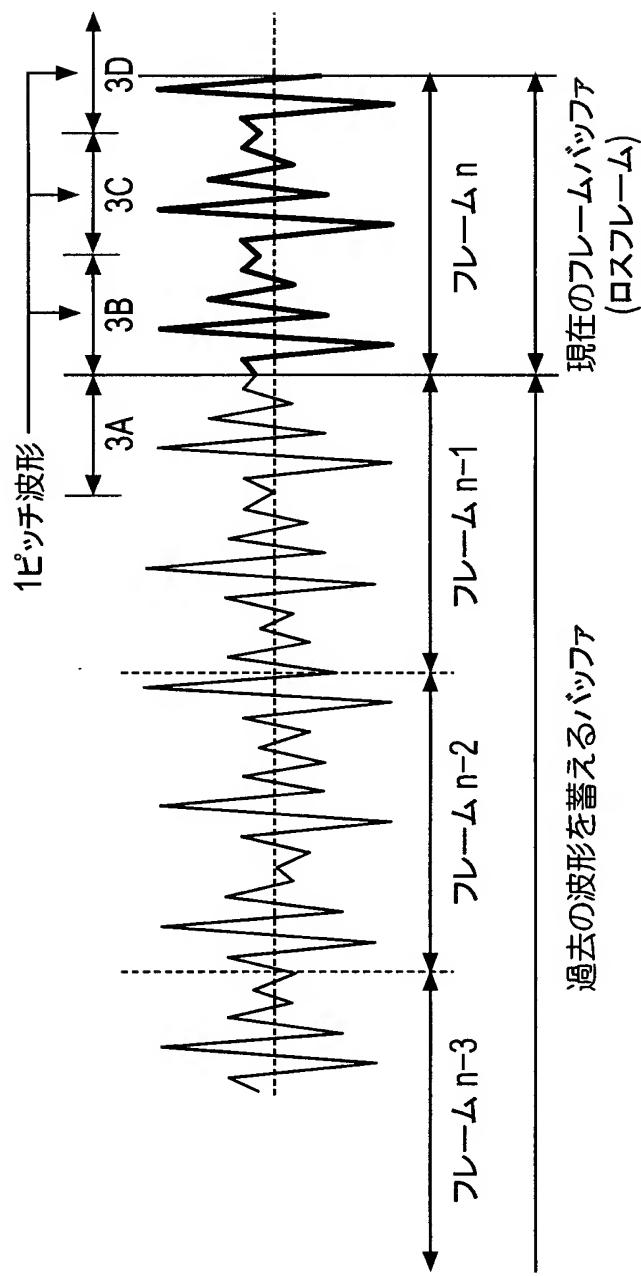


図4

【図 5】

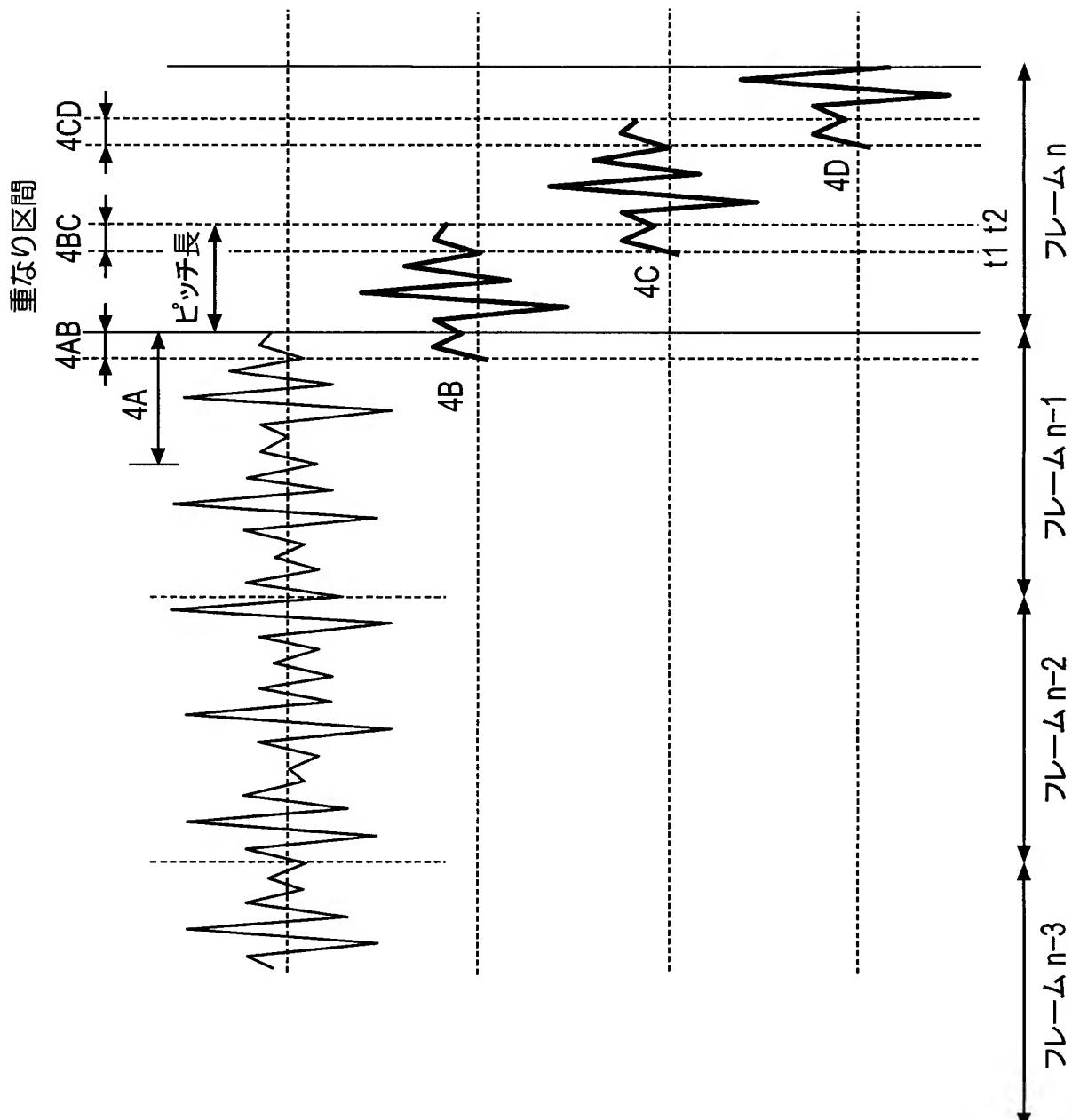


図5

【図 6】

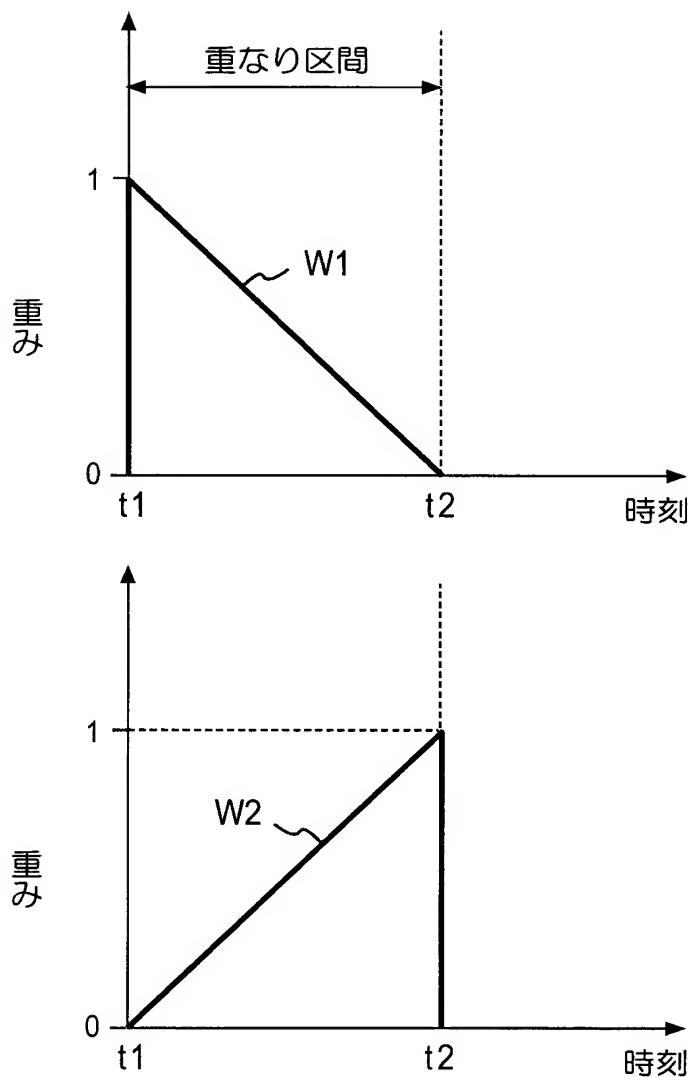


図6

【図 7】

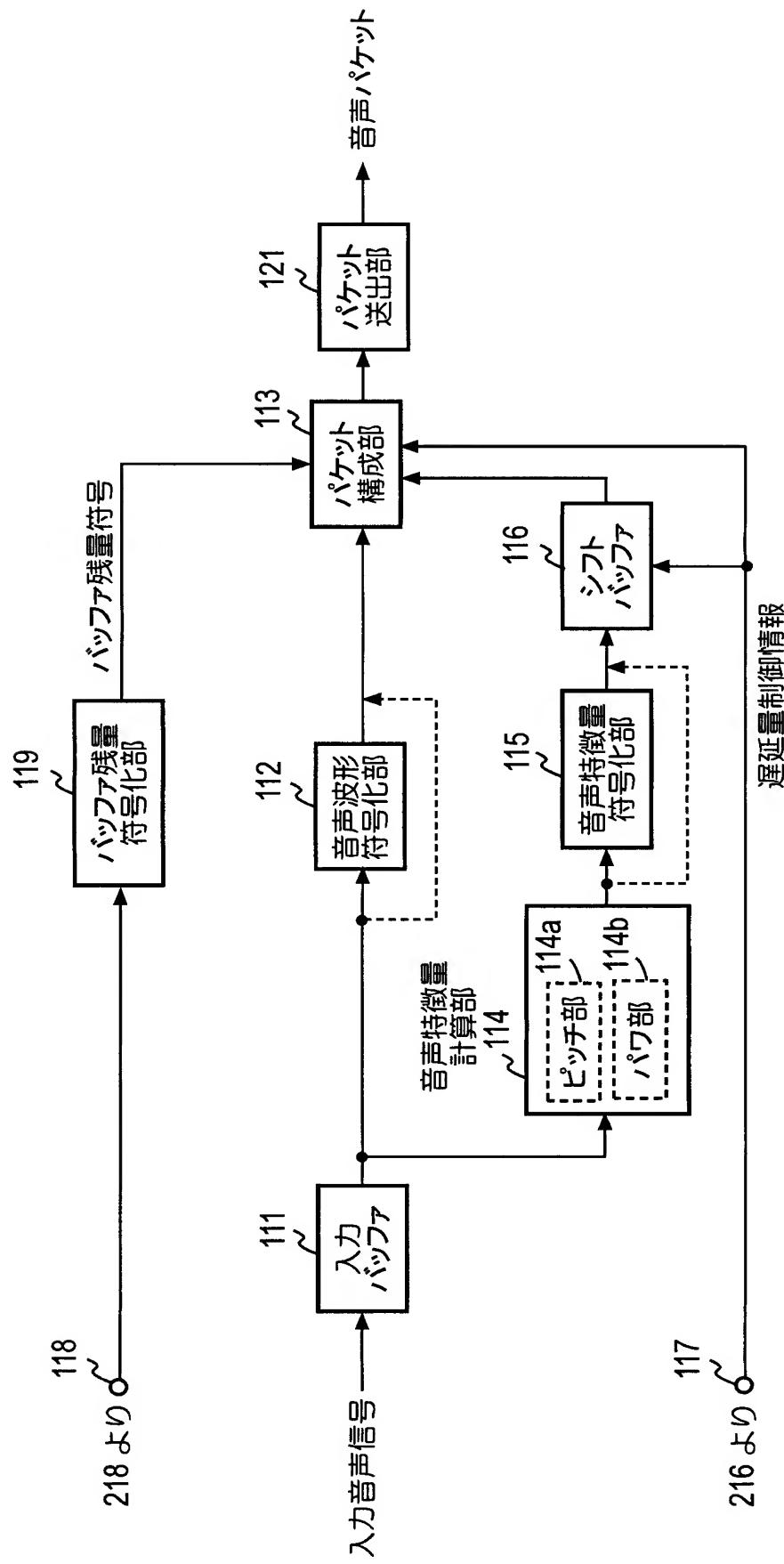
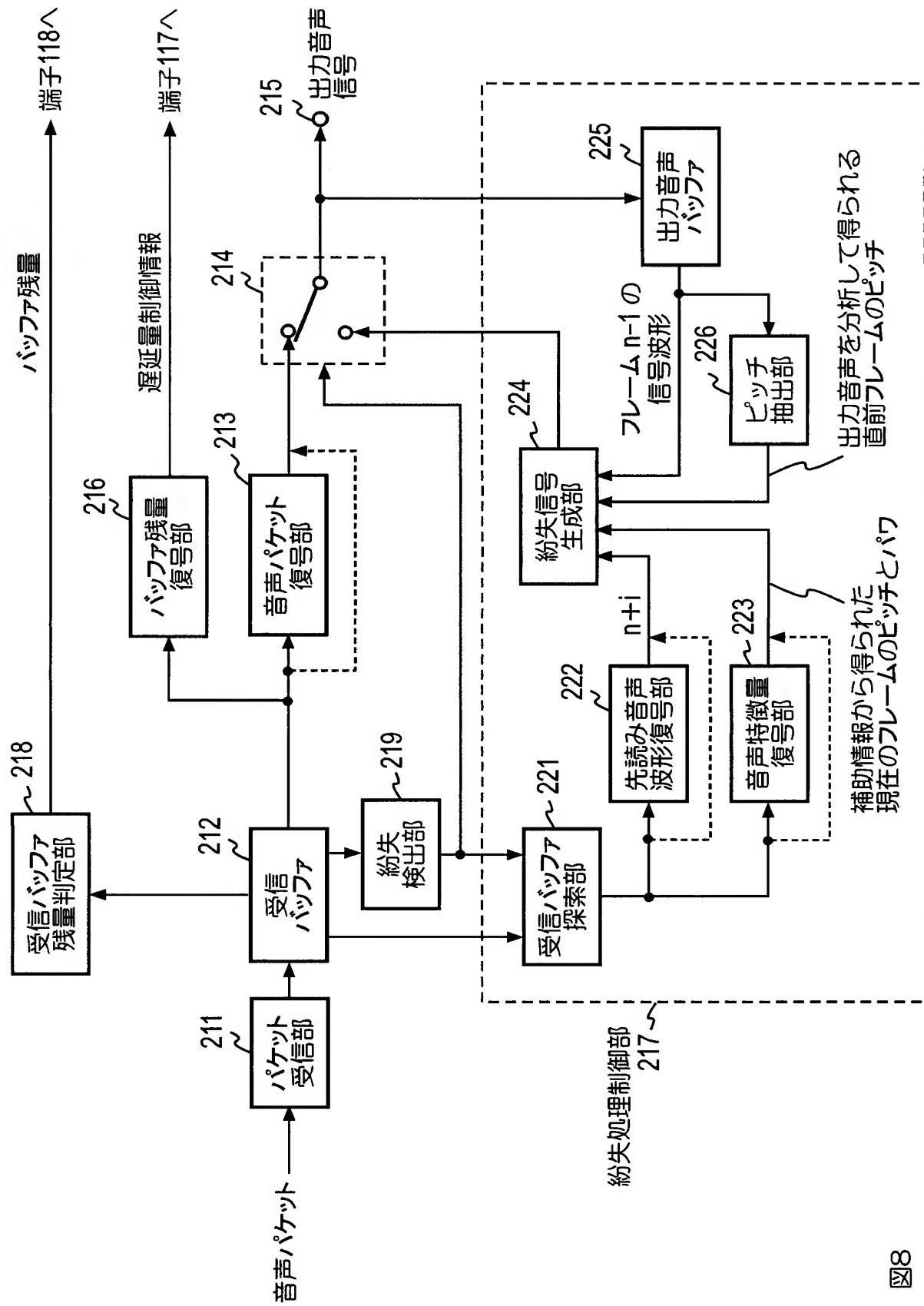


図7

【図8】



88

【図 9】

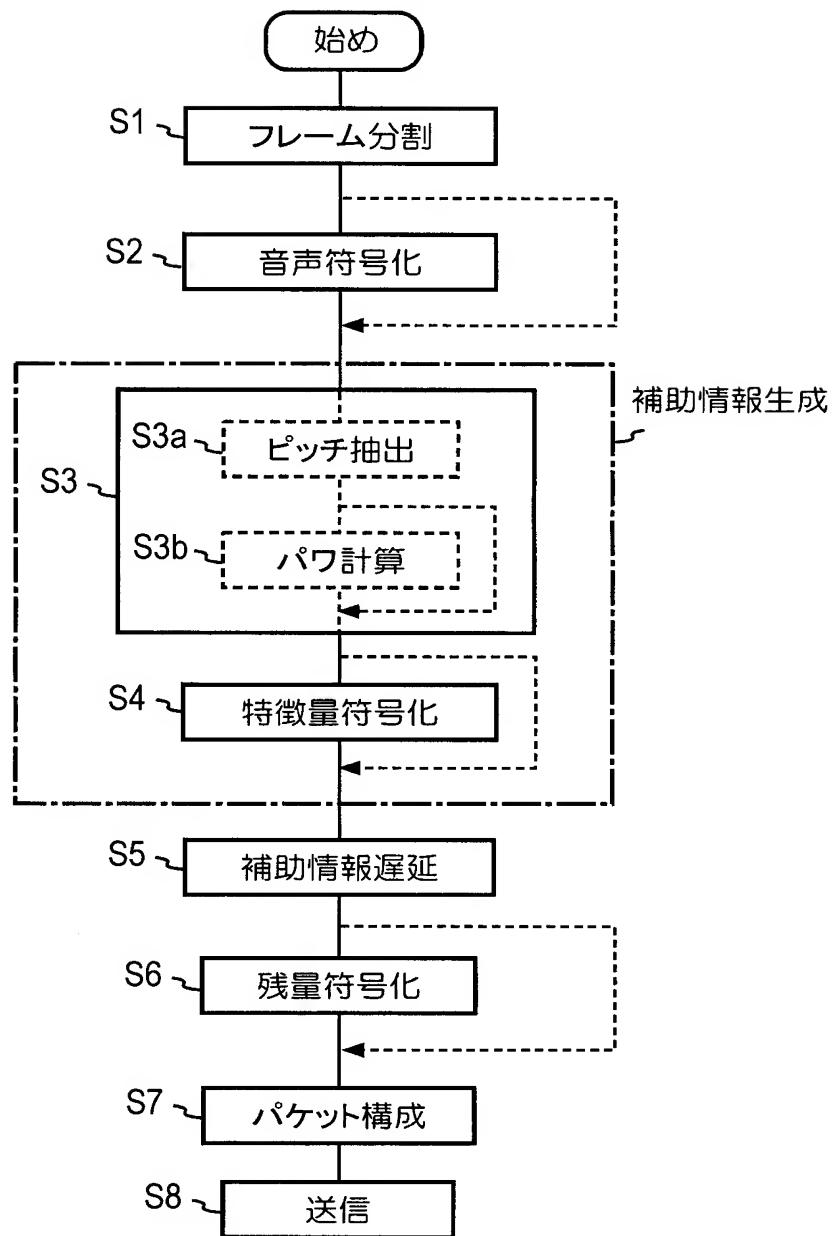


図9

【図10】

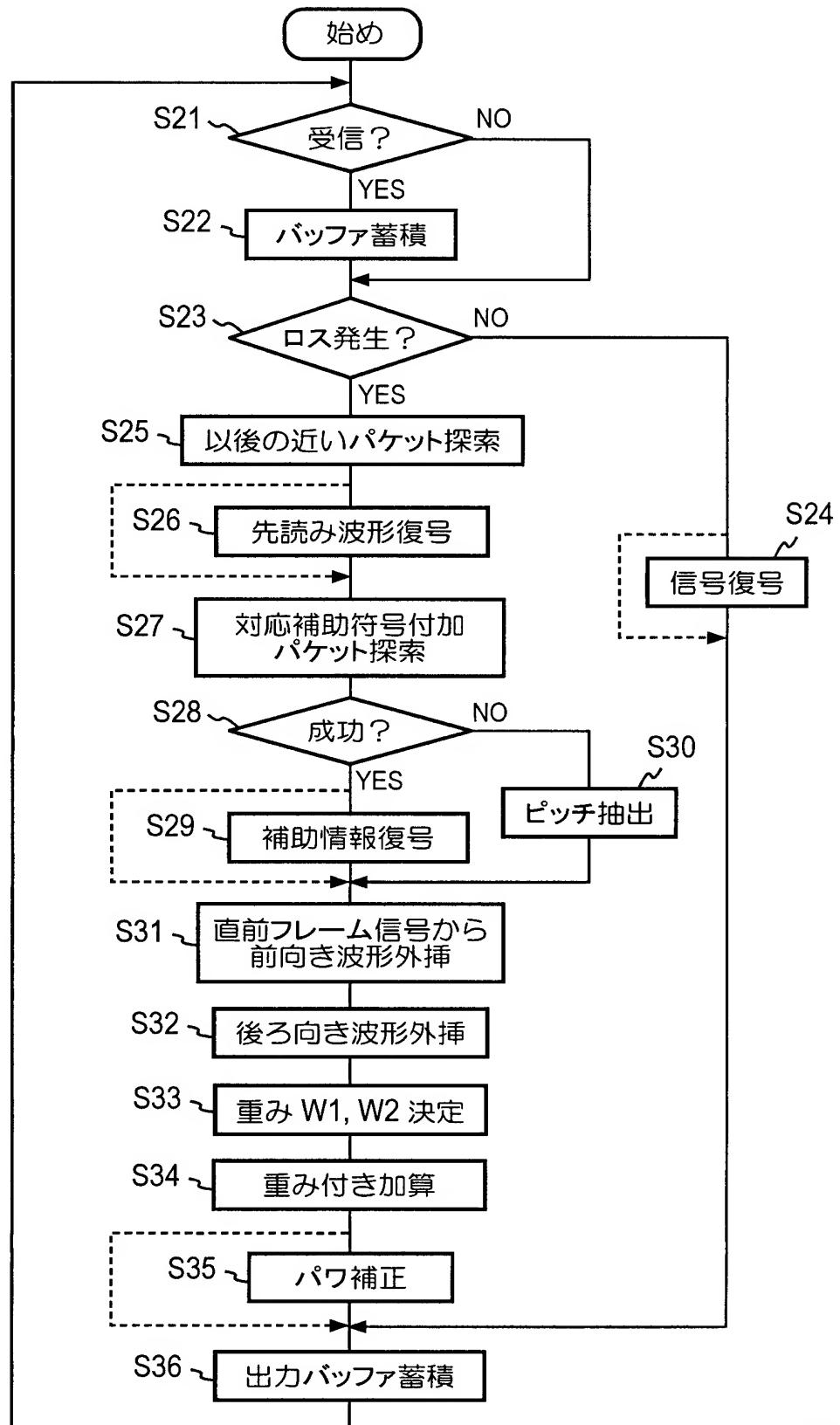


図10

【図 1-1】

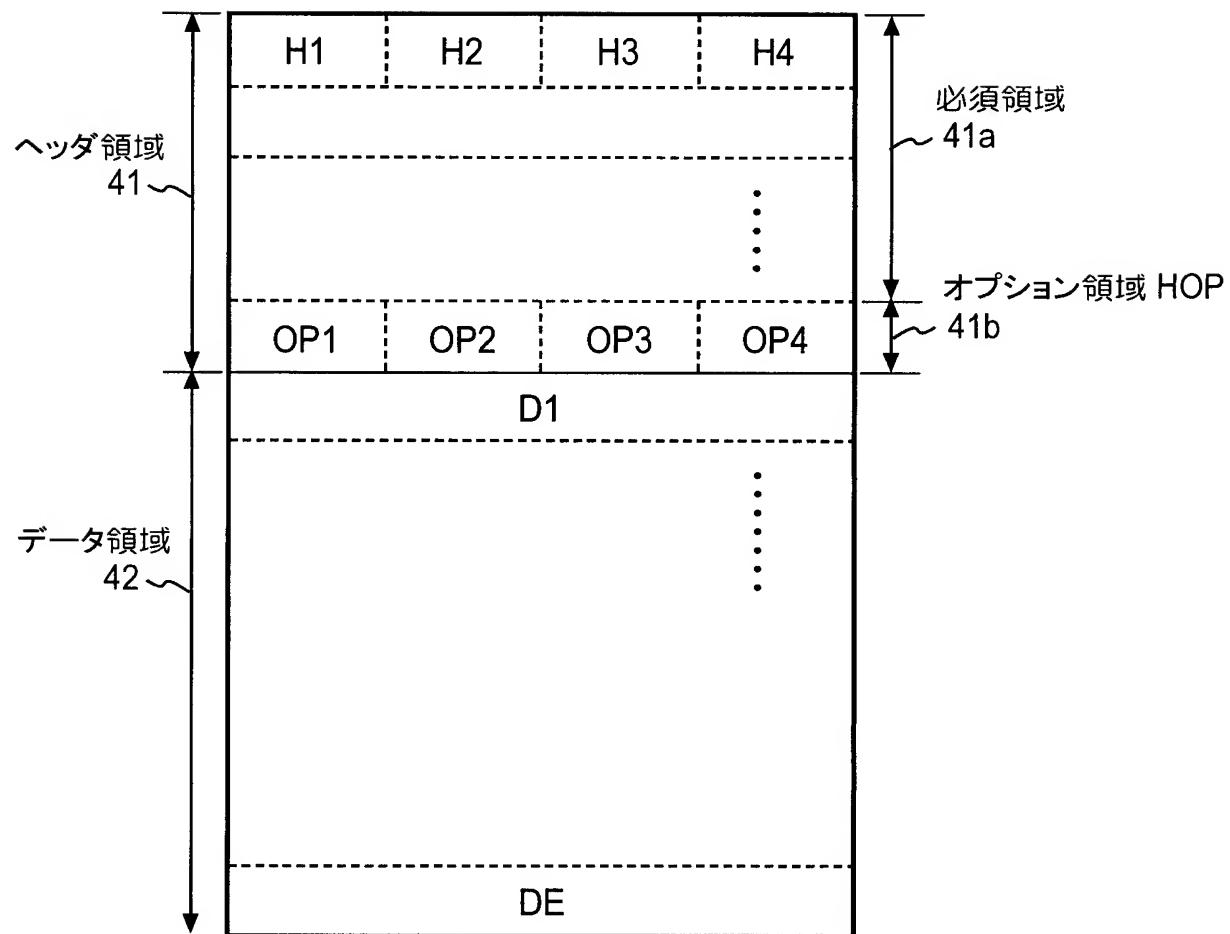


図1-1

【図 12】

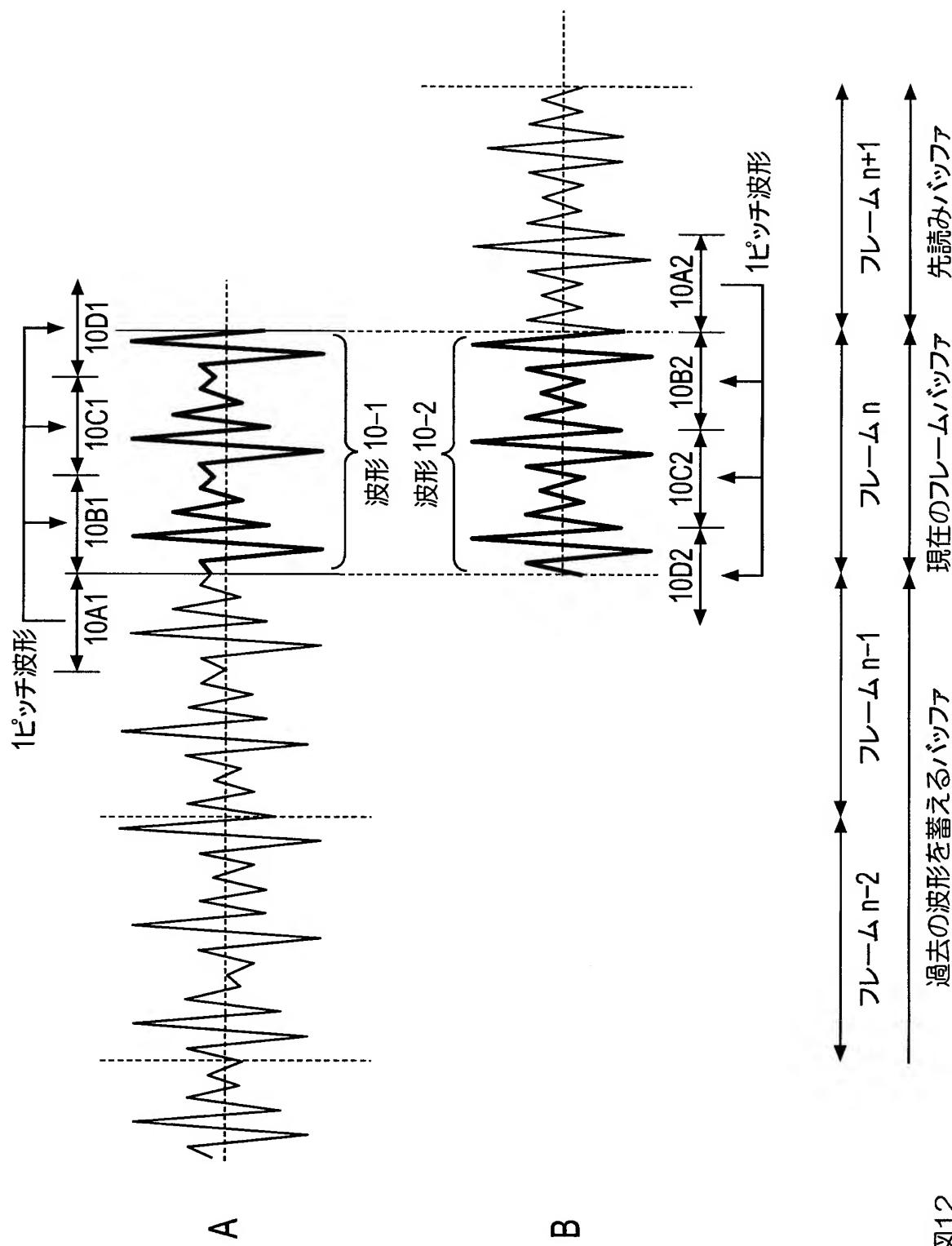


図12

【図 13】

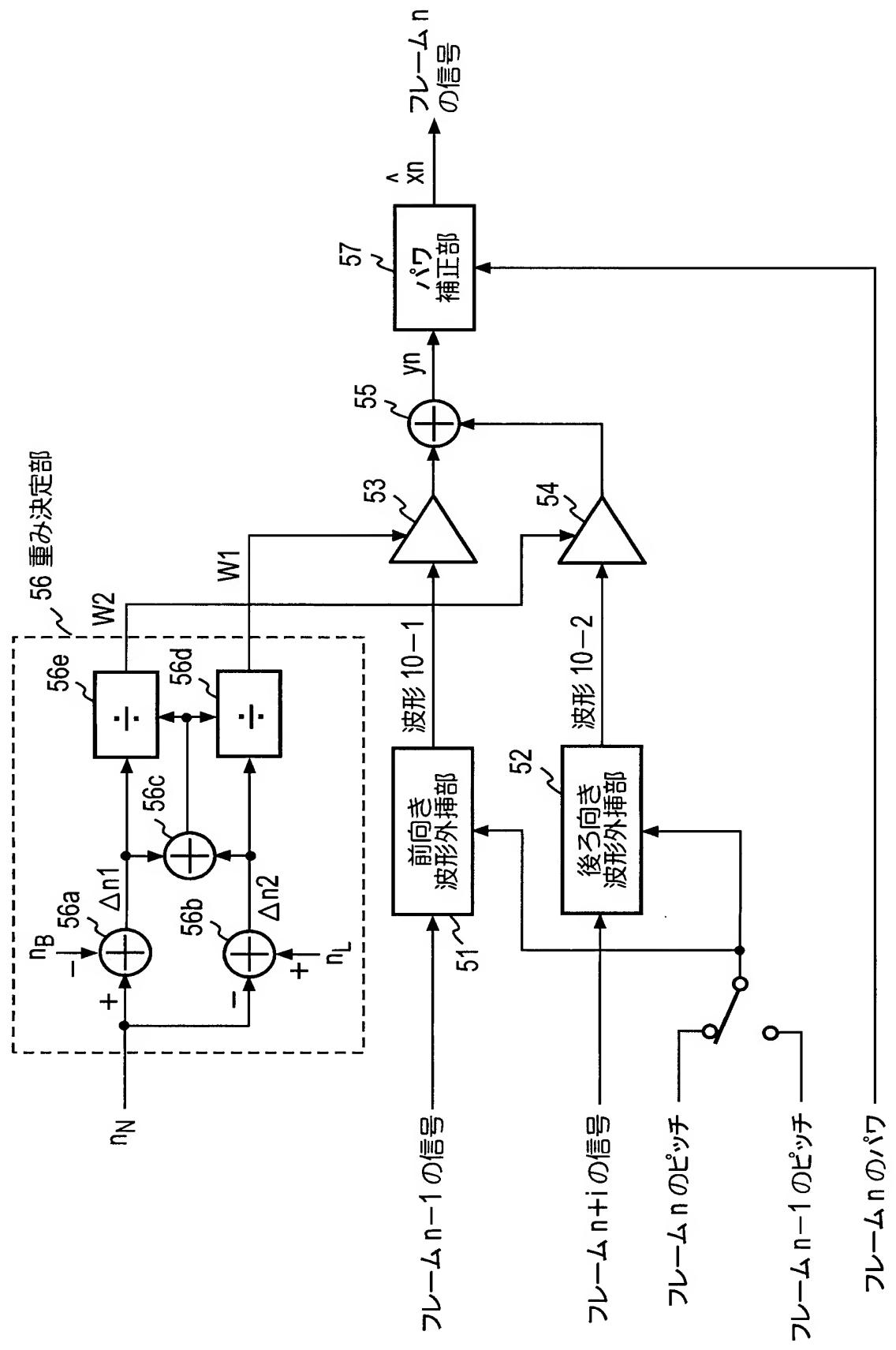


図13

【図 14】

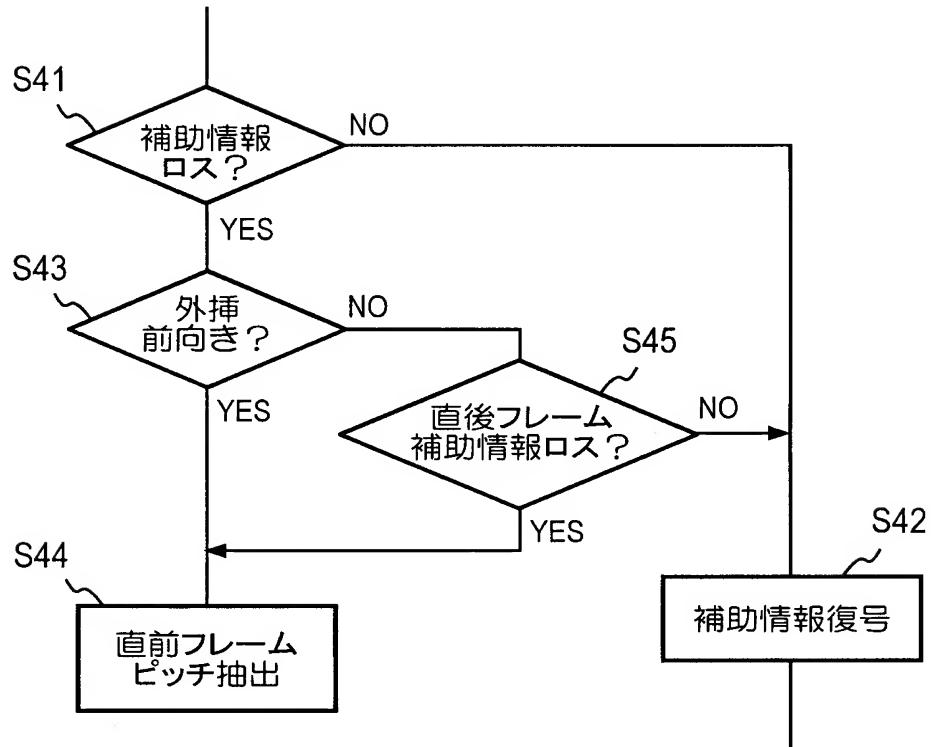


図14

【図 15】

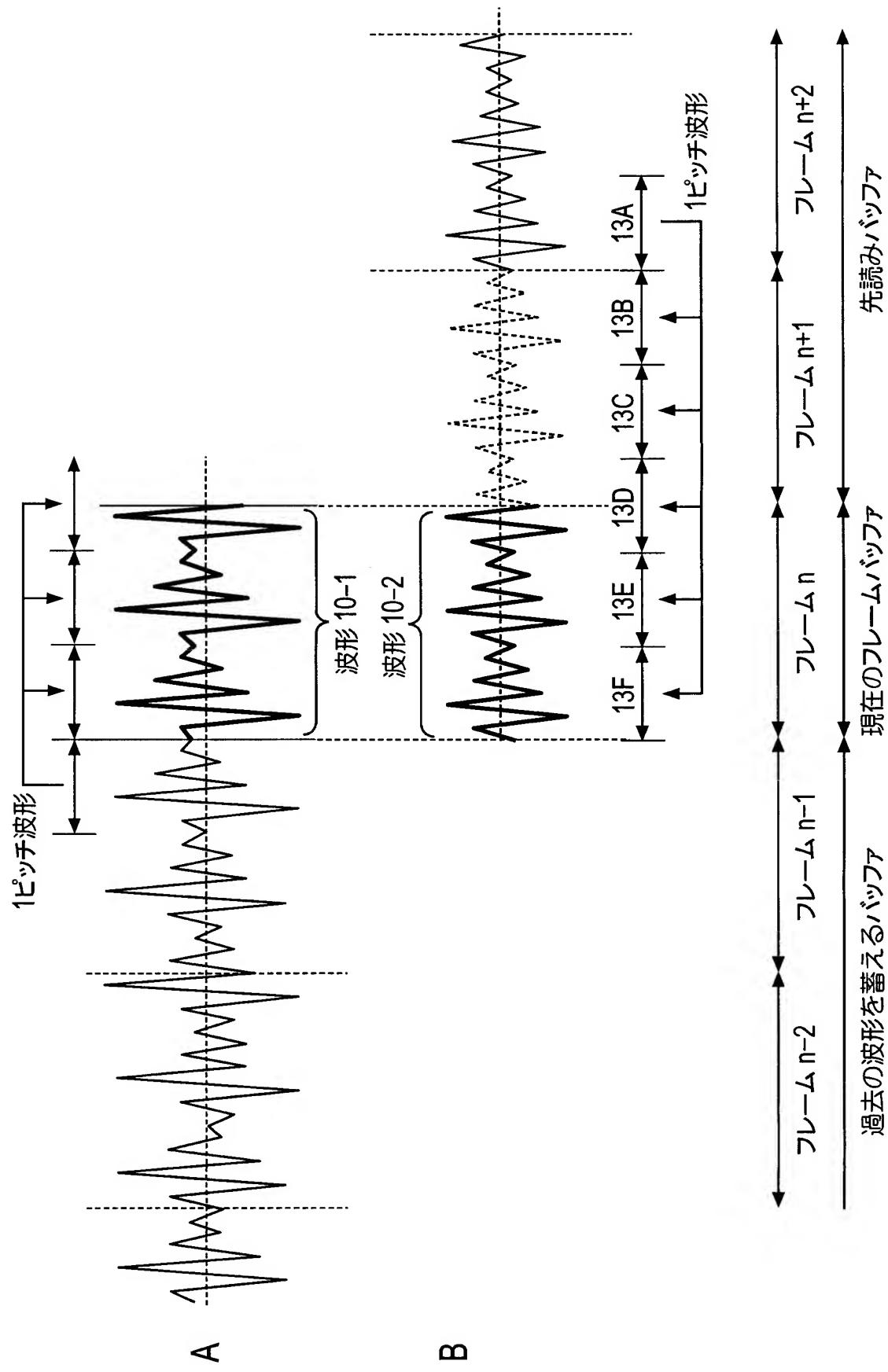


図15

【図 16】

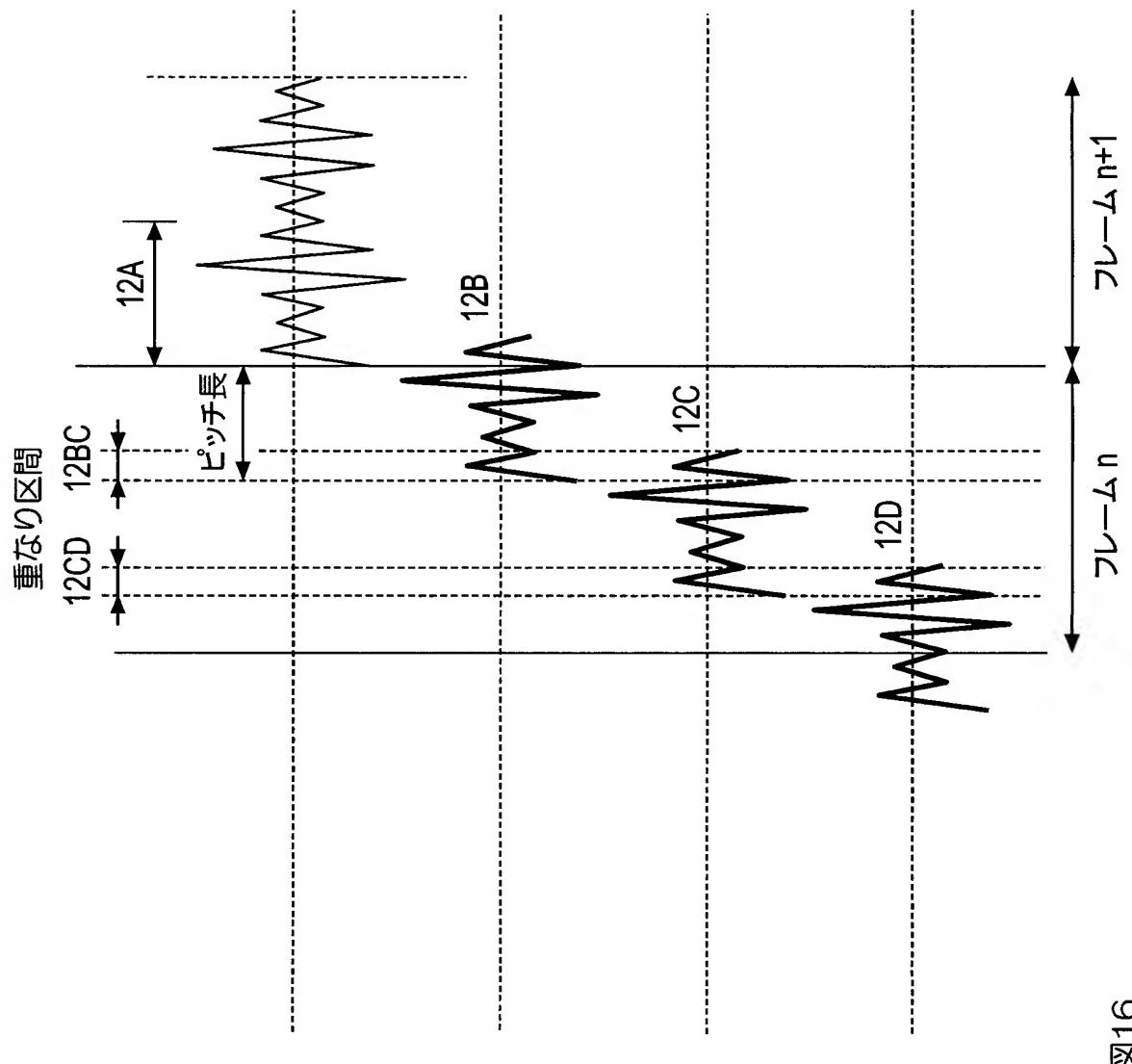


図16

【書類名】要約書

【要約】

【課題】 子音と母音の境界付近のフレームのパケットの紛失やバースト的パケット紛失に対して再生音声の品質を向上させる。

【解決手段】 送信側で、フレームnの音声信号のピッチ長及びパワを、フレームnのパケットより後のフレームのパケットに付加して送信し、受信側でフレームnのパケットが紛失すると、フレームnの信号のピッチ長、パワが付加されたパケットからピッチ長、パワを取り出し、フレームn-1及びフレームn+1の各波形から取り出したピッチ長の波形10A1及び10A2をそれぞれ取り出し、前者で10B1, 10C1, 10D1と前向き外挿し、後者で10B2, 10C2, 10D2と後ろ向き外挿し、これら両外挿波形を重み付き加算して、フレームnの信号とする。

【選択図】 図12

出願人履歴

0 0 0 0 0 4 2 2 6

19990715

住所変更

5 9 1 0 2 9 2 8 6

東京都千代田区大手町二丁目3番1号

日本電信電話株式会社